

What is claimed is:

- 13
- 09705069 140200
1. A method for classifying signals comprising:
dividing an input signal into blocks having a predetermined time length;
extracting one or more than one characteristic quantities of a signal attribute from the signal of each block; and
classifying the signal of each block into a category according to the characteristic quantities thereof.
 2. The method for classifying signals according to claim 1, wherein said signal of each block is classified into any of the categories formed on the basis of types of signal sources.
 3. The method for classifying signals according to claim 1, wherein said signal of each block is classified into any of the categories formed on the basis of types structures that signals may have and do not depend on the types of signal sources.
 4. The method for classifying signals according to claim 2, wherein said input signal is an audio signal; and
the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound

sources.

5. The method for classifying signals according to claim 3, wherein said input signal is an audio signal; and

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where only a sound related to a single sound source exists in the block, a double sound source structure where sounds related respectively to two sound sources exist in the block, a sound source change structure where a sound source including silence is switched only for once in the block, a multiple sound source change structure where a plurality of sound sources are switched simultaneously in the block, a sound source partial change structure where part of a plurality of sound sources are switched in the block and an extra structure pattern where none of the above patterns is applicable and are used for categorical classification based on the structures.

6. The method for classifying signals according to claim 1, wherein one or more than one of the average and variances of the signal power in the block, the average and variances of the power of a band-pass signal of the signal in the block, the average and variances of the spread of the spectrogram of the signal in the block, the average and variances of the pitch frequency of the

signal in the block, the average and variances of the degree of harmonic structurization of the signal in the block, the average and variances of the residue signal of linear predictive analysis of the signal in the block and the average and variances of the pitch gain of the residue signal of linear predictive analysis of the signal in the block are used as said characteristic quantities.

7. The method for classifying signals according to claim 6, wherein said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

8. The method for classifying signals according to claim 1, wherein a vector quantization technique is used as method for the categorical classification.

9. An apparatus for classifying signals comprising:

a blocking means for dividing an input signal into blocks having a predetermined time length;

a feature extracting means for extracting one or more than one characteristic quantities of a signal attribute from the signal of each block; and

a categorical classifying means for classifying the signal of each block

into a category according to the characteristic quantities thereof.

10. The apparatus for classifying signals according to claim 9, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of signal sources.

11. The apparatus for classifying signals according to claim 9, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of structures that signals may have and do not depend on the types of signal sources.

12. The apparatus for classifying signals according to claim 10, wherein said input signal is an audio signal; and

the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound sources.

13. The apparatus for classifying signals according to claim 11, wherein

said input signal is an audio signal; and

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio

signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where only a sound related to a single sound source exists in the block, a double sound source structure where sounds related respectively to two sound sources exist in the block, a sound source change structure where a sound source including silence is switched only for once in the block, a multiple sound source change structure where a plurality of sound sources are switched simultaneously in the block, a sound source partial change structure where part of a plurality of sound sources are switched in the block and an extra structure pattern where none of the above patterns is applicable and are used for categorical classification based on the structures.

14. The apparatus for classifying signals according to claim 9, wherein said feature extracting means uses one or more than one of the average and variances of the signal power in the block, the average and variances of the power of a band-pass signal of the signal in the block, the average and variances of the spread of the spectrogram of the signal in the block, the average and variances of the pitch frequency of the signal in the block, the average and variances of the degree of harmonic structurization of the signal in the block, the average and variances of the residue signal of linear predictive analysis of the signal in the block and the average and variances of the pitch gain of the residue signal of linear predictive analysis of the signal in the block

as said characteristic quantities.

15. The apparatus for classifying signals according to claim 14, wherein said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

16. The apparatus for classifying signals according to claim 9, wherein said categorical classifying means uses a vector quantization technique as method for the categorical classification.

17. A method for generating descriptors comprising:
dividing an input signal into blocks having a predetermined time length;
extracting one or more than one characteristic quantities of a signal attribute from the signal of each block;

classifying the signal of each block into a category according to the characteristic quantities thereof; and

generating a descriptor for the signal according to the category of classification thereof.

18. The method for generating descriptors according to claim 17, wherein said signal of each block is classified into any of the categories formed

on the basis of types of signal sources.

19. The method for generating descriptors according to claim 17, wherein said signal of each block is classified into any of the categories formed on the basis of types of structures that signals may have and do not depend on the types of signal sources.

20. The method for generating descriptors according to claim 18, wherein

said input signal is an audio signal; and

the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound sources.

21. The method for generating descriptors according to claim 19, wherein

said input signal is an audio signal;

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where

AB

1. The first step is to identify the problem or goal.

ound related to a single sound source exists in the
 urce structure where sounds related respectively to t
 the block, a sound source change structure where
 g silence is switched only for once in the block,
 change structure where a plurality of sound sour
 eously in the block, a sound source partial change str
 ality of sound sources are switched in the bock and
 where none of the above patterns is applicable
 cal classification based on the structures; and
 descriptor is generating according to the categor
 the structures.

. The method for generating descriptors accor
 one or more than one of the average and variances o
 ock, the average and variances of the power of a b
 al in the block, the average and variances of t
 ram of the signal in the block, the average and vari
 y of the signal in the block, the average and varia
 onic structurization of the signal in the block,
 s of the residue signal of linear predictive analysis
 d the average and variances of the pitch gain of th
 redictive analysis of the signal in the block

characteristic quantities.

23. The method for generating descriptors according to claim 22, wherein

said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

24. The method for generating descriptors according to claim 17, wherein a vector quantization technique is used as method for the categorical classification.

25. An apparatus for generating descriptors comprising:

a blocking means for dividing an input signal into blocks having a predetermined time length;

a feature extracting means for extracting one or more than one characteristic quantities of a signal attribute from the signal of each block;

a categorical classifying means for classifying the signal of each block into a category according to the characteristic quantities thereof; and

a descriptor generating means for generating a descriptor for the signal according to the category of classification thereof.

26. The apparatus for generating descriptors according to claim 25, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of signal sources.

27. The apparatus for generating descriptors according to claim 25, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of structures that signals may have and do not depend on the types of signal sources.

28. The apparatus for generating descriptors according to claim 26, wherein

said input signal is an audio signal; and

the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound sources.

29. The apparatus for generating descriptors according to claim 27, wherein

said input signal is an audio signal;

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio

AB 09705059 410300

signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where only a sound related to a single sound source exists in the block, a double sound source structure where sounds related respectively to two sound sources exist in the block, a sound source change structure where a sound source including silence is switched only for once in the block, a multiple sound source change structure where a plurality of sound sources are switched simultaneously in the block, a sound source partial change structure where part of a plurality of sound sources are switched in the block and an extra structure pattern where none of the above patterns is applicable and are used for categorical classification based on the structures; and

said descriptor generating means generates a descriptor according to the categorical classification based on the structures.

30. The apparatus for generating descriptors according to claim 25, wherein said feature extracting means uses one or more than one of the average and variances of the signal power in the block, the average and variances of the power of a band-pass signal of the signal in the block, the average and variances of the spread of the spectrogram of the signal in the block, the average and variances of the pitch frequency of the signal in the block, the average and variances of the degree of harmonic structurization of the signal in the block, the average and variances of the residue signal of linear predictive

analysis of the signal in the block and the average and variances of the pitch gain of the residue signal of linear predictive analysis of the signal in the block as said characteristic quantities.

31. The apparatus for generating descriptors according to claim 30, wherein

AB
said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

69705059 110200
said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

32. The apparatus for generating descriptors according to claim 25, wherein said categorical classifying means uses a vector quantization technique as method for the categorical classification.

33. A method for retrieving signals comprising:
dividing an input signal into blocks having a predetermined time length;
extracting one or more than one characteristic quantities of a signal attribute from the signal of each block;
classifying the signal of each block into a category according to the characteristic quantities thereof; and
retrieving the signal according to the result of categorical classification

or by using a descriptor generated according to the result of categorical classification.

34. The method for retrieving signals according to claim 33, wherein said signal of each block is classified into any of the categories formed on the basis of types of signal sources.

35. The method for retrieving signals according to claim 33, wherein said signal of each block is classified into any of the categories formed on the basis of types of structures that signals may have and do not depend on the types of signal sources.

36. The method for retrieving signals according to claim 34, wherein said input signal is an audio signal;

the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound sources; and

a signal is retrieved by using the descriptor reflecting or corresponding to the result of said categorical classification based on the sound sources.

37. The method for retrieving signals according to claim 35, wherein said input signal is an audio signal;

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where only a sound related to a single sound source exists in the block, a double sound source structure where sounds related respectively to two sound sources exist in the block, a sound source change structure where a sound source including silence is switched only for once in the block, a multiple sound source change structure where a plurality of sound sources are switched simultaneously in the block, a sound source partial change structure where part of a plurality of sound sources are switched in the block and an extra structure pattern where none of the above patterns is applicable and are used for categorical classification based on the structures; and

a signal is retrieved by using the descriptor reflecting or corresponding to the result of said categorical classification based on the structure.

38. The method for retrieving signals according to claim 33, wherein one or more than one of the average and variances of the signal power in the block, the average and variances of the power of a band-pass signal of the signal in the block, the average and variances of the spread of the spectrogram of the signal in the block, the average and variances of the pitch frequency of the signal in the block, the average and variances of the degree of harmonic

structurization of the signal in the block, the average and variances of the residue signal of linear predictive analysis of the signal in the block and the average and variances of the pitch gain of the residue signal of linear predictive analysis of the signal in the block are used as said characteristic quantities.

39. The method for retrieving signals according to claim 38, wherein said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

40. The method for retrieving signals according to claim 33, wherein a vector quantization technique is used as method for the categorical classification.

41. The method for retrieving signals according to claim 33, wherein points of changes of the signal are detected by using the descriptor reflecting or corresponding to the result of said categorical classification.

42. A apparatus for retrieving signals comprising:

a blocking means for dividing an input signal into blocks having a predetermined time length;

a feature extracting means for extracting one or more than one

characteristic quantities of a signal attribute from the signal of each block;

a categorical classifying means for classifying the signal of each block into a category according to the characteristic quantities thereof; and

a signal retrieving means for retrieving the signal according to the result of categorical classification or by using a descriptor generated according to the result of categorical classification.

43. The apparatus for retrieving signals according to claim 42, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of signal sources.

44. The apparatus for retrieving signals according to claim 42, wherein said categorical classifying means classifies said signal of each block into any of the categories formed on the basis of types of structures that signals may have and do not depend on the types of signal sources.

45. The apparatus for retrieving signals according to claim 43, wherein said input signal is an audio signal;

the categories formed on the basis of signal sources for classifying the audio signal of each block include one or more than one of silence, voice, male voice, female voice, music, vocal music, instrumental music, noise, striking sound, environmental sound, sound of hustle and bustle, clapping sound and cheering sound and are used for categorical classification based on the sound sources; and

said signal retrieving means retrieves a signal by using the descriptor reflecting or corresponding to the result of said categorical classification based on the sound sources.

46. The apparatus for retrieving signals according to claim 44, wherein said input signal is an audio signal;

the categories formed on the basis of structures that signals may have and do not depend on the types of signal sources for classifying the audio signal of each block include one or more than one of a silence structure where no significant sound exists in the block, a single sound source structure where only a sound related to a single sound source exists in the block, a double sound source structure where sounds related respectively to two sound sources exist in the block, a sound source change structure where a sound source including silence is switched only for once in the block, a multiple sound source change structure where a plurality of sound sources are switched simultaneously in the block, a sound source partial change structure where part of a plurality of sound sources are switched in the block and an extra structure pattern where none of the above patterns is applicable and are used for categorical classification based on the structures; and

said signal retrieving means retrieves a signal by using the descriptor reflecting or corresponding to the result of said categorical classification based on the structure.

47. The apparatus for retrieving signals according to claim 42, wherein said feature extracting means uses one or more than one of the average and variances of the signal power in the block, the average and variances of the power of a band-pass signal of the signal in the block, the average and variances of the spread of the spectrogram of the signal in the block, the average and variances of the pitch frequency of the signal in the block, the average and variances of the degree of harmonic structurization of the signal in the block, the average and variances of the residue signal of linear predictive analysis of the signal in the block and the average and variances of the pitch gain of the residue signal of linear predictive analysis of the signal in the block as said characteristic quantities.

48. The apparatus for retrieving signals according to claim 47, wherein said average of the degree of harmonic structurization is the temporal average of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies; and

said variances of the degree of harmonic structurization is the temporal standard deviation of the ratio of the energy of the sound component of integer times of the pitch frequency to the energy of all the frequencies.

49. The apparatus for retrieving signals according to claim 42, wherein said categorical classifying means uses a vector quantization technique as method for the categorical classification.

50. The apparatus for retrieving signals according to claim 42, wherein said signal retrieving means detects points of changes of the signal by using the descriptor reflecting or corresponding to the result of said categorical classification.

AB

00205069 40200